

## Meeting Summary

# Workshop on Ensemble Forecasting in the Short to Medium Range

Thomas M. Hamill

*NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado*

James A. Hansen

*Department of Earth, Atmospheric, and Planetary Sciences  
Massachusetts Institute of Technology, Cambridge, Massachusetts*

Steven L. Mullen

*Department of Atmospheric Sciences  
University of Arizona, Tucson, Arizona*

Chris Snyder

*National Center for Atmospheric Research, Boulder, Colorado*

Submitted to *Bulletin of the American Meteorological Society*

25 January 2004

### Corresponding author address

Dr. Thomas M. Hamill

NOAA-CIRES Climate Diagnostics Center

R / CDC 1

325 Broadway

Boulder, CO 80305 – 3328

Ph: 1 (303) 497-3060 Fax: 1 (303) 497-7013

e-mail: [tom.hamill@noaa.gov](mailto:tom.hamill@noaa.gov)

## 1. Background

The most radical change to numerical weather prediction (NWP) during the last decade has been the operational implementation of ensemble forecast methods. Rather than applying all of the available computational resources to a single, highest-resolution simulation, some portion are allocated to an ensemble with somewhat reduced-resolution. Ensemble forecasts offer a way of filtering the predictable from the unpredictable through averaging – the features that are consistent among ensemble members are preserved, while those that are inconsistent are reduced in amplitude. Perhaps more important, the ensemble itself, as a sample from possible forecast outcomes, can be used to estimate the forecast uncertainty and the likely structure of forecast errors.

The first operational implementations consisted of a trial of a few to a few dozen global ensemble forecasts initialized from a variety of slightly different initial conditions. Ten years later, ensemble forecasting is arguably the mainstay of medium-range NWP and is being integrated into short-range NWP. Ensembles of forecasts now commonly are based on multiple models or model configurations as well as multiple initial conditions. A wide variety of techniques have been developed for extracting information from ensembles and presenting the voluminous data to forecasters and sophisticated users. Techniques have also been developed for using the information on forecast uncertainty in the data assimilation process.

A previous workshop on ensemble forecasting was held in September, 1999, at the National Center for Atmospheric Research. A summary of this meeting was discussed in Hamill et al (2000). Four years later, we convened again in Val-Morin, Quebec from 18-20 September 2003 to assess the state of the art of ensemble forecasting,

to discuss the most substantial problems, and to map research directions for the coming years.

Before discussing the workshop itself, let's consider the research progress made through the last four years. Clearly, ensemble forecast research is growing rapidly. Over 100 peer-reviewed articles on ensemble forecasting have been published since the last workshop in 1999, a larger number than in all the years before 1999. Further, this publication count does not include articles on ensemble-forecast related topics such as adaptive observations, basic predictability research, ensemble forecasting of climate, or applications to oceanography or other geosciences.

One very active area of research involves methods for initializing ensemble forecasts. Ideally, the initial perturbations would be sampled from the probability distribution of plausible analysis states (e.g., Ehrendorfer and Tribbia 1997, Hamill et al. 2003). In practice, however, operational centers attempt to approximate this ideal using various strategies. The European Centre for Medium-Range Weather Forecasts (ECMWF) uses a “singular vector” technique, whereby the perturbations are specifically designed to grow as rapidly as possible over the first few days of the forecast (e.g., Molteni et al. 1996). The National Centers for Environmental Prediction (NCEP) uses a “breeding” technique (Toth and Kalnay, 1993, 1997) that produces perturbations reflecting where errors have grown quickly in the recent past. The Canadian Meteorological Centre (CMC) uses a “perturbed observation” approach (Houtekamer et al. 1996), where parallel cycles of first-guess forecasts are updated to distinct sets of perturbed observations, producing an ensemble of analyses.

Research continues into methods for initializing ensemble forecasts. The dynamical and statistical properties of singular-vector ensembles have been widely studied (Barkmeijer et al. 1999, Reynolds and Palmer, 1999, Reynolds and Rosmond 2003, Hamill et al. 2003, Gelaro et al. 2002), examined for application to short-range, limited-area modeling (Ehrendorfer et al. 1999, Frogner and Iversen 2001, 2002, Hersbach et al. 2000, 2003), and for generating perturbations in the tropics (Barkmeijer et al., 2001, Puri et al. 2001, Zhang and Krishnamurti, 1999). The characteristics of bred vectors have been considered (Errico and Langland 1999, Toth et al. 1999, Patil et al. 2001). New approaches to generating perturbations such as the Ensemble Transform Kalman Filter (ETKF; Bishop et al. 2001, Wang and Bishop 2003, 2004) and the “indistinguishable states” approach (Judd and Smith 2001) have been developed that more closely attempt to sample the distribution of analysis states. And a range of comparisons of the various perturbation methods has been performed (Trevisan et al. 2001, Cheung 2001, Miller et al. 2002, Hamill et al. 2003, Wang and Bishop 2003).

A different approach to generating initial conditions is *ensemble data assimilation*. Rather than generating perturbations in some manner around a control analysis, ensemble data assimilation methods conduct an ensemble of parallel data assimilation cycles, like the existing operational scheme at CMC. Further, the ensemble of first-guess forecasts are used to estimate flow-dependent forecast error statistics during the assimilation, which may improve the quality of the set of analyses. A large volume of literature has appeared during the last four years, including Anderson (2001, 2003), Etherton and Bishop (2003), Hamill et al. (2001), Hansen and Smith (2001), Hansen (2002), Heemink et al. (2001), Houtekamer and Mitchell (2001), Keppenne (2000),

Keppenne and Rienecker (2002), Lermusiaux and Robinson (1999), Miller et al. (1999), Mitchell and Houtekamer (2002), Ott et al. (2003), Pham (2001), Reichle et al. (2002 a,b), Snyder and Zhang (2003), Verlaan and Heemink (2001), Whitaker and Hamill (2001), and Zhang et al. (2003). Several reviews papers are also available, including Tippett et al. (2003), Evensen (2004), and Lorenc (2004). To date, most of the work has been in simplified models with synthetic data, so the applicability to actual atmospheric data assimilation has yet to be demonstrated. However, some recent real-data studies suggest that ensemble-based assimilation schemes may be competitive with or superior to existing methods (Houtekamer et al. 2004, Whitaker et al. 2004, Dowell et al. 2004).

Since ensembles of forecasts are used to assess the uncertainty in the weather prediction, an ensemble ought to produce a realistically diverse set of simulations. Too often, the ensembles of forecasts unduly resemble each other. This may be due to several factors. Perhaps the ensemble was started from a set of initial conditions that did not sample the distribution of plausible analysis states. And without question, the imperfections in the model itself can bias the ensemble. Forecast models may have systematic errors, so that on average the ensemble mean may be consistently too warm or too dry. Additionally, because ensemble forecasts are conducted at a finite resolution, the full spectrum of atmospheric motions and their interactions are not properly represented.

How problematic is model error? Orrell et al. (2001) and Orrell (2002) suggest that the uncertainty due to model error is dramatically larger than that contributed by initial condition deficiencies through chaos. A wide variety of other studies, on the other hand, suggest the growth of errors due to initial condition deficiencies is generally larger

(e.g., Simmons and Hollingsworth 2002, Fig. 6). Moreover, definitively attributing forecast errors to initial condition deficiencies or model errors is problematic. For example, an infinitesimal model error incurred during the beginning of a simulation results in a difference between forecast states that can thereafter increase in magnitude due to chaotic processes. Regardless of whether model error is predominant or merely important, the ensemble forecast community recognizes that the uncertainty of forecasts due to model imperfections needs to be addressed in ensemble forecasting and ensemble data assimilation.

A variety of approaches to account for forecast-model error have been proposed, with various levels of success. One possibility that has been explored is to conduct an ensemble of forecasts using a variety of models (e.g., Evans et al. 2000, Krishnamurti et al. 2000a, Mylne et al. 2002, Richardson 2001a, Wandishin et al. 2001) or a variety of parameterizations or model configurations (e.g., Grell and Devenyi 2002, Gritmit and Mass 2002). Another possibility is to statistically adjust the ensemble forecasts (Du et al. 2000, Krishnamurti 2000ab, Hamill et al. 2004, Legg et al. 2002, Roulston and Smith 2003). Another possible alternative is to formulate the NWP model stochastically, so that the uncertainty in the time tendency is simulated through the integration of random noise (Palmer 2001, Sardeshmukh et al. 2001). Several relevant studies of model error have also been performed in the last few years in simple systems (e.g. Smith et al. 1999, Vannitsem and Toth 2002) and more complex ones (Tribbia and Baumhefner 2003).

Much of what the community has learned about the characteristics of ensemble forecasts has come through synoptic evaluations and verification studies (e.g., Atger 2001, Bright and Mullen 2002, Buizza et al. 2000, 2001, Buizza and Chessa 2002, Buizza

and Hollingsworth 2001, Colucci et al. 1999, Ebert 2001, Mass et al. 2003, Molteni and Buizza 1999, Pelly and Hoskins 2003, Roulston et al. 2003, Sanders et al. 2000, Toth et al. 2001, Watson and Colucci 2002). Several studies have shown the beneficial effects of higher resolution in ensemble forecasts, which generally increases their spread (Buizza et al. 2003, Elmore et al. 2002ab, 2003, Gritti and Mass 2002, Hersbach et al. 2000, Mullen and Buizza 2002, Szunyogh and Toth 2002), the aspects of multi-model ensembles (Evans et al. 2000, Krishnamurti et al. 2000a, Richardson 2001a, Wandishin et al. 2001), applications to tropical weather (Cheung 2001, Krishnamurti et al. 2000b, Mackey and Krishnamurti 2001, Puri et al. 2001), and to medium-range weather forecasting (Hamill et al. 2004).

Ensembles are naturally suited to producing probabilistic forecasts, and so methods for evaluating them are required. Tools for assessing the economic value have become a standard way of evaluating ensemble forecasts (Buizza 2001, Mylne 2002, Palmer 2002, Richardson 2001b, Thornes and Stephenson 2001, Wilks 2001, Zhu et al. 2002). Other evaluation methods include alternatives to clustering methods (Atger 1999), techniques for assessing the nonlinearity of forecasts (Gilmour et al. 2001), rank histograms (Hamill 2001), a multidimensional extension known as the “minimum spanning tree” (Smith 2000, Wilks 2003), information theory diagnostics (Roulston and Smith 2002), and other methods (e.g., Lalaurette 2003, Stephenson and Doblas-Reyes 2000, Toth et al. 2001, Wei and Toth 2003).

The rest of the summary consists of a report on the research and discussion presented at the workshop. There were three sessions, a session on the treatment of model error in ensemble forecasts, the verification and use of ensemble forecasts, and

methods for initializing ensemble forecasts, including ensemble data assimilation.

Approximately 60 scientists attended this workshop from the U.S., Canada, and Europe.

## **2. Summary of ensemble workshop sessions**

### *a. Session 1: The treatment of model error*

The NWP models used in ensemble forecasting are not perfect; commonly this manifests itself in a biased ensemble with too little spread. The first session of the workshop discussed ways to treat model errors in ensemble forecasts. Methods discussed included multi-model ensembles, stochastic parameterizations, stochastic-dynamic prediction, and other methods.

More than half of the presentations focused on model-error issues in limited-area models. These less expensive, limited-area ensembles can be run using initial and lateral boundary conditions supplied by operational global NWP ensemble forecasts. Many presentations examined the impact of running multi-models/multi-parameters/multi-parameterization ensembles. Plans were presented for no fewer than three operational, regional multi-model ensemble forecast systems for Europe, as well as descriptions of two operational multi-model/multi-analysis ensemble forecast systems in North America. Each effort to account for model error was shown to provide positive benefits, but so far, these results are based on techniques that are empirically rather than theoretically justified. To what extent is it necessary to take a multi-model/multi-parameter/multi-parameterization approach to ensemble forecasting? Does perturbing a model's

parameters provide something that perturbing physics tendencies cannot, or vice versa? The field has not yet matured enough to provide answers to these questions.

Several presentations demonstrated that there are methods to account for model error that can increase ensemble forecast spread. But is the extra spread “good” spread? Increasing a forecast ensemble’s variance does not necessarily mean that the eigenvectors of the error covariances estimated from the ensemble are oriented in the correct directions, or that the ensemble is located in the correct region of state space. One presentation dealing with empirically boosting the ensemble variance in Fourier space demonstrated that while the magnitude of the ensemble spread can be corrected, significant problems remain with the phase of the forecast uncertainty. Not surprisingly, this indicates that model error impacts the structure of forecast uncertainty as well as its magnitude. Another presentation clearly demonstrated the benefits of multi-model and multi-physics ensembles for quantitative precipitation forecasting. This result suggests that the multi-model/multi-physics approach is capable of making important structural changes to the forecast uncertainty.

Will it be possible to disentangle the forecast error due to model error from the forecast error due to initial-condition error and boundary-condition error? Many examinations of model error assumed that the ensemble of initial conditions properly sample the analysis-error uncertainty and the boundary conditions provide suitable large-scale variability. If these assumptions are wrong, then errors in the forecast ensemble cannot be definitively attributed to one problem or the other. This highlights the continuing importance of research on methods for generating initial conditions (section

2c). If initial and boundary conditions can be properly designed, then subsequent forecast error can more realistically be attributed to model error.

*b. Session 2: Verification and use of ensemble forecasts*

This session included talks on new methods for ensemble verification, use of ensembles and economic value studies, methods of communicating information from ensembles, post-processing methods, regional experiments, and predictability studies.

Several talks and ensuing discussion concentrated on the validation of ensemble forecasts. A goal of ensemble forecasting is to produce as sharp as possible a forecast PDF that is still reliable. Accurately validating ensemble forecasts is complicated by three factors, the observational error, the finiteness ( $N$ ) of the ensemble size, and the finiteness of the number of cases that are evaluated ( $M$ ). The imperfect nature of the validation data, whether from observations or gridded analyses, has historically not been considered in probabilistic verification studies. A presentation at the workshop showed that it is theoretically preferable and feasible to build a verification system that treats the validation data as a random variable. The other two complications are sampling issues. Even if an ensemble draws from the appropriate PDF, probability estimates based on ensemble relative frequency will become more inaccurate as  $N$  decreases. Furthermore, assessing the reliability of these probabilistic forecasts is difficult without  $M$  being very large, especially when assessing the reliability of rare events. Given these inherent limitations, it was discussed that ensemble validation ought to regularly include information on the uncertainty of the results such as confidence intervals.

Several case studies illustrated the strengths and weaknesses of current ensemble forecast systems. One study showed that none of the operational analyses from the major forecast centers correctly initialized a developing storm in the eastern Pacific. None of the subsequent numerical predictions correctly forecast the landfall of the subsequent storm. This illustrates that there are still weaknesses in data assimilation and methods for initializing ensemble forecasts. Another presentation showed that short-range ensembles for two U.S. East-Coast snowstorms varied in their performance. In one case the ensemble provided valuable additional information to the deterministic run, indicating the possibility of a snowstorm, while in another case, every ensemble member missed the snowstorm.

Presentations by researchers at ECMWF, NCEP, and several other institutions indicated that progress is being made to improve the operational ensemble forecasts. The current ECMWF ensemble was shown to have positive spread-skill relationships and on average provides probabilistic forecasts with greater economic value than their corresponding deterministic forecast. However, ECMWF has noted several severe-weather cases where the ensemble forecasts did not encompass the actual severe weather event. Accordingly, they plan to increase the resolution of their ensemble, update their method of determining initial conditions to include “moist singular vectors” (Barkmeijer et al. 2001), and make several other changes. NCEP is exploring a different method of generating initial perturbations, using the ensemble transform Kalman filter (Wang and Bishop 2003). They are also increasing the resolution of their global ensemble, adding more members, and exploring methods of treating model error and ensemble-based data assimilation methods.

An ensemble modeling system that produces perfectly calibrated forecasts directly from the raw ensemble counts is likely to remain an elusive goal over the next several years. Hence, post-processing of output from ensemble forecast systems was broadly discussed. All viable processing schemes appear capable of significantly improving reliability, but commonly, the post-processing techniques cannot sharpen the forecasts while improving reliability. A range of post-processing techniques were discussed, from model-output statistics approaches (Hamill et al. 2004) to “Bayesian model averaging” (Kass and Raftery 1995). The length of ensemble output needed for training was a subject of debate. Simple bias removal in error-prone models might require only a short training period of a few weeks to provide some benefit, but for other applications such as medium-range forecasting, years of training data may be needed to achieve optimal results.

*c. Session 3: Methods of generating initial conditions and ensemble data assimilation*

The session covered methods of generating initial conditions for ensemble forecasts, including research in ensemble-based data assimilation methodologies. The session began with talks that focused on the generation of initial conditions. These included the generalization of the breeding technique to systems with multiple time scales and the ETKF as an alternative approach to singular vectors and breeding.

The majority of presentations dealt with ensemble data assimilation, consistent with its emergence as an area of substantial research activity over the last four years. (Indeed, even the generation of ensemble perturbations via the ETKF is closely related to

ensemble data assimilation.) Numerous ensemble assimilation methods were presented and discussed, but almost all of these were similar in their broad outline, particularly in their use of: (1) an ensemble of nonlinear forecasts to estimate statistics of short-range (background) forecast of both model variables and observed variables; (2) an analysis step much like the standard Kalman filter, and (3) an explicit assumption that forecast error correlations are significant only over limited distances. The details of the various methods differ, however. Many of the differences reside in the analysis step. Some methods process each observation sequentially while others compute analysis increments in one step based on all the observations valid at a given time (or at least using batches of observations in a given region). Some methods also include numerical approximations in the calculation of the analysis, for example to improve parallelization in distributed-memory computing architectures. Some methods were stochastic, treating the observations as random variables, while others were deterministic.

The methods presented at the workshop also differed in how they accounted for the uncertainty produced by errors in forecast model and by sampling errors in the assimilation algorithm itself. In all cases, this additional uncertainty was accounted for by increasing the spread of the forecast ensemble through empirical/ad hoc methods. Both additive approaches, in which noise drawn from a known distribution is added to the forecast ensemble, and multiplicative approaches, in which the deviations from the ensemble mean are increased by a scalar factor, are being explored. Bias correction techniques are also being explored (see section 2a).

The presentations also considered a variety of applications for ensemble data assimilation schemes. Several of the presentations, along with much of the existing

literature in the atmospheric sciences, concentrated on global- and synoptic-scale flows (with an eye to algorithms suitable for operational numerical weather prediction). In addition, participants in the workshop presented studies focusing on the tropical ocean, atmospheric meso- and convective-scale flows, limited-area models, and soil moisture analyses.

Several broad conclusions and issues emerged from the discussion at the workshop:

1. Research in a variety of application has now shown that ensemble assimilation schemes work well in experiments with simulated observations and a perfect forecast model. Early results with real observations are, perhaps not surprisingly, more mixed. Further experience with real observations will clearly be necessary in order to have a better grasp of the capabilities of ensemble data assimilation and to compare the advantages and disadvantages of different schemes.

2. As with ensemble forecasting in general, there is a need to improve the way in which errors in the forecast model are accounted for in ensemble assimilation schemes. Many existing schemes use very simple techniques in this regard; because they attempt direct and time-evolving estimates of forecast-error covariances, their performance may well be significantly limited by their treatment of model error.

3. Additional tools are needed for diagnosing the performance of ensemble assimilation schemes. While diagnostics may be (and have been) borrowed from ensemble forecasting or data assimilation, ensemble data assimilation poses specific questions that have not been extensively explored elsewhere, such as the quality of estimated covariances.

4. The role of nonlinear and non-Gaussian effects in data assimilation and their treatment in ensemble assimilation schemes remains an open question. One approach put forward at the workshop was to develop models for forecast error statistics using variables or coordinates for which the errors are more nearly Gaussian, such as considering displacement errors of coherent features.

### **3. Summary**

A workshop on ensemble forecasting was held in Val-Morin, Quebec, from 18-20 September 2003. The meeting discussed recent and current research on ensemble forecasts. The workshop consisted of a session on methods for initializing ensemble forecasts, including ensemble data assimilation, a session on model error, and a session on the use and interpretation of ensemble forecasts. A theme running through all three sessions was the effects of model error on ensemble forecasts. Model error biases the ensemble and results in a lack of spread. Methods are being formulated for diagnosing and treating model error directly in ensemble forecasts, estimating its effects in ensemble data assimilation, and correcting it through post-processing to produce calibrated probability forecasts.

## References

- Anderson, J. L., 2001: An ensemble adjustment filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.
- Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634-642.
- Atger, F., 1999: Tubing: an alternative to clustering for the classification of ensemble forecasts. *Wea. Forecasting*, **14**, 741-757.
- Atger, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlin. Proc. in Geophys.*, **8**, 401-417.
- Atger, F., 2002: Spatial and interannual variability of the reliability of ensemble-based probabilistic forecasts: consequences for calibration. *Mon. Wea. Rev.*, **131**, 1509-1523.
- Barkmeijer, J., Buizza, R., and Palmer, T.N., 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2333-2351.
- Barkmeijer, J., Buizza, R., Palmer, T.N., Puri, K., and Mahfouf, J.-F., 2001: Tropical singular vectors computed with linearized diabatic physics. *Quart. J. Roy. Meteor. Soc.*, **127**, 685-708.
- Bishop, C. H., B. J. Etherton, and S. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part 1: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420-436.

- Bright, D. R., Mullen, S. L. 2002: Short-range ensemble forecasts of precipitation during the southwest monsoon. *Wea. Forecasting*, **17**, 1080-1100.
- Buizza, R., J. Barkmeijer, T. N. Palmer, and D. Richardson, 2000: Current status and future developments of the ECMWF ensemble prediction system. *Meteor. Appl.*, **7**, 163-175.
- Buizza, R., A. Hollingsworth, F. LaLaurette, and A. Ghelli, 2001: Probability precipitation prediction using the EMCWF ensemble prediction system. *Wea. Forecasting*, **14**, 168-189.
- Buizza, R. 2001: Accuracy and potential economic value of categorical and probabilistic forecasts of discrete events. *Mon. Wea. Rev.*, **129**, 2329-2345.
- Buizza, R., and A. Hollingsworth, 2001: Storm prediction over Europe using the ECMWF ensemble prediction system. *Meteor. Appl.*, **9**, 289-306.
- Buizza, R., and Chessa, P. 2002: Prediction of the U.S. storm of 24-26 January 2000 with the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **130**, 1531-1551.
- Buizza, R., D. S. Richardson, and T. N. Palmer, 2003: Benefits of increased resolution in the ECMWF ensemble system and comparisons with poor-man's ensembles. *Quart. J. Roy. Meteor. Soc.*, **129**, 1269-1288.
- Cheung, K. W., 2001: Ensemble forecasting of tropical cyclone motion: comparison between regional bred modes and random perturbations. *Meteor. Atmos. Phys.*, **78**, 23-34.
- Colucci, S. J., Baumhefner, D. P., and Konrad, C. E. 1999: Numerical prediction of a cold-air outbreak: a case study with ensemble forecasts. *Mon. Wea. Rev.*, **127**, 1538-1550.

- Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and thermodynamic retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: ensemble Kalman filter experiments. *Mon. Wea. Rev.*, submitted.
- Du, J., Mullen, S. L., and Sanders, F. 2000: Removal of distortion error from an ensemble forecast. *Mon. Wea. Rev.*, **128**, 3347-3351.
- Ebert, E. E., 2001: Ability of a poor-man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461-2480.
- Ehrendorfer, M., and J. J. Tribbia, 1997: Optimal prediction of forecast-error covariances through singular vectors. *J. Atmos. Sci.*, **54**, 286-313.
- Ehrendorfer, M., Errico, R. M., and Raeder, K. D., 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.*, **56**, 1627-1648.
- Elmore, Kim. L., Stensrud, D. J., and Crawford, K. C. 2002: Ensemble cloud model applications to forecasting thunderstorms. *J. Appl. Meteor.*, **41**, 363-383.
- Elmore, K. L., Stensrud, D. J., Crawford, K. C. 2002: Explicit cloud-scale models for operational forecasts: a note of caution. *Wea. Forecasting*, **17**, 873-884.
- Elmore, K. L., Weiss, S. J., Banacos, P. C. 2003: Operational ensemble cloud model forecasts: some preliminary results. *Wea. Forecasting*, **18**, 953-964.
- Errico, R. M., and Langland, R., 1999: Notes on the appropriateness of "bred modes" for generating initial perturbations used in ensemble prediction. *Tellus*, **51A**, 431-441.
- Etherton, B. J., and C. H. Bishop, 2003: Resilience of hybrid ensemble/3D-Var analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.*, in review. Available from [betherto@rsmas.miami.edu](mailto:betherto@rsmas.miami.edu).

- Evans, R. E., Harrison, M.S.J., Graham, R.J. Mylne, K.R., 2000: Joint medium-range ensembles from the Met. Office and ECMWF systems. *Mon. Wea. Rev.*, **128**, 3104-3127.
- Evensen, G. 2004: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, submitted.
- Frogner, I.-L., and Iversen, T., 2001: Targeted ensemble prediction for northern Europe and parts of the North Atlantic ocean. *Tellus*, **53**, 35-55.
- Frogner, I.-L., and T. Iversen, 2002: High-resolution limited-area ensemble predictions based on low-resolution targeted singular vectors. *Quart. J. Roy. Meteor. Soc.*, **128**, 1321-1341.
- Gelaro, R., T. Rosmond, and R. Daley, 2002: Singular vector calculations with an analysis-error covariance metric. *Mon. Wea. Rev.*, **130**, 1166-1186.
- Gilmour, I., L. A. Smith, and R. Buizza, 2001: Linear regime duration: is 24 hours a long time in synoptic weather forecasting? *J. Atmos. Sci.*, **58**, 3525-3539.
- Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Letters*, **29**, 38-1 - 38-4.
- Grimit, E. P., and C. F. Mass, 2002: Initial results of a mesoscale short-range ensemble forecasting system over the Pacific northwest. *Wea. Forecasting*, **17**, 192-205.
- Hamill, T. M., S. L. Mullen, C. Snyder, Z. Toth, and D. P. Baumhefner, 2000: Ensemble forecasting in the short to medium range. Report from a workshop. *Bull. Amer. Meteor. Soc.*, **81**, 2653-2664.

- Hamill, T. M., C. Snyder, and J. S. Whitaker, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776-2790.
- Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550-560.
- Hamill, T. M., C. Snyder, and J. S. Whitaker, 2003: Ensemble forecasts and the properties of flow-dependent analysis-error covariance singular vectors. *Mon. Wea. Rev.*, **131**, 1741-1757.
- Hamill, T. M., J. S. Whitaker, and X. Wei, 2004: Ensemble re-forecasting: improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, in press. Available from <http://www.cdc.noaa.gov/~hamill>.
- Hansen, J. A., and L. A. Smith, 2001: Probabilistic noise reduction. *Tellus*, **53A**, 585-598.
- Hansen, J. A., 2002: Accounting for model error in ensemble-based state estimation and forecasting. *Mon. Wea. Rev.*, **130**, 2373-2391.
- Heemink, A. W., M. Verlaan, and A. J. Segers, 2001: Variance-reduced ensemble Kalman filtering. *Mon. Wea. Rev.*, **129**, 1718-1728.
- Hersbach, H., R. Mureau, J. D. Opsteegh, and J. Barkmeijer, 2000: A short-range to early-medium-range ensemble prediction system for the European area. *Mon. Wea. Rev.*, **128**, 3501-3519.
- Hersbach, H., R. Mureau, J. D. Opsteegh, and J. Barkmeijer, 2003: Development of a targeted ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **129**, 2027-2048.

- Houtekamer, P. L., L. Lefaivre, and J. Derome, 1996: The RPN ensemble prediction system. Proc. ECMWF Seminar on Predictability. Vol II, Reading, United Kingdom, 121-146. Available from [www.ecmwf.int](http://www.ecmwf.int).
- Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123-137.
- Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2004: Atmospheric data assimilation with the ensemble Kalman filter: results with real observations. *Mon. Wea. Rev.*, submitted. Available from [peter.Houtekamer@ec.gc.ca](mailto:peter.Houtekamer@ec.gc.ca).
- Judd, K., and L. A. Smith, 2001: Indistinguishable states I: the perfect model scenario. *Physica D*, **151**, 125-141.
- Kass, R. E., and A. E. Raftery, 1995: Bayes factors. *Journal of the American Statistical Association*, **90**, 773-795.
- Keppenne, C. L., 2000: Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Mon. Wea. Rev.*, **128**, 1971-1981.
- Keppenne, C. L., and M. M. Rienecker, 2002: Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.*, **130**, 2951-2965.
- Krishnamurti, T. N., Kishtawal, C. M., Zhang, Z., LaRow, T., Bachiochi, D., and Williford, C. E., 2000a: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196-4216.

- Krishnamurti, T. N., Kishtawal, C. M., Shin, D. W., Williford, C. E., 2000b: Improving tropical precipitation forecasts from a multianalysis superensemble. *J. Climate*, **13**, 4217-4227.
- LaLaurette, F., 2003: Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Quart. J. Roy. Meteor. Soc.*, **129**, 3037-3057.
- Legg, T. P., Mylne, K.R., and Woolcock, C., 2002: Use of medium-range ensembles at the Met Office I: PREVIN - a system for the production of probabilistic forecast information from the ECMWF EPS. *Meteorol. Appl.*, **9**, 225-271.
- Lermusiaux, P. F. J., and A. R. Robinson, 1999: Data assimilation via error subspace statistical estimation. *Mon. Wea. Rev.*, **127**, 1385-1407.
- Lorenc, A. C, 2004: The potential of the ensemble Kalman filter for NWP - a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, in press.
- Mackey, B. P., and Krishnamurti, T. N., 2001: Ensemble forecast of a typhoon flood event. *Wea. Forecasting*, **16**, 399-415.
- Mass, C. F., and coauthors, 2003: Regional environmental prediction over the Pacific Northwest. *Bull. Amer. Meteor. Soc.* **84**, 1353-1366.
- Miller, R. N., E. F. Carter, and S. T. Blue, 1999: Data assimilation into nonlinear stochastic models. *Tellus*, **51A**, 167-194.
- Miller, R. N., Ehret, L. L. 2002: Ensemble generation for models of multimodal systems. *Mon. Wea. Rev.*, **130**, 2313-2333.
- Mitchell, H. L., and P. L. Houtekamer, 2002 : Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791-2808.

- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliaxis, 1996: The ECMWF ensemble prediction system: methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73-119.
- Molteni, F., and Buizza, R. 1999: Validation of the ECMWF ensemble prediction system using empirical orthogonal functions. *Mon. Wea. Rev.*, **127**, 2346-2358.
- Mullen, S. L., and R. Buizza, 2002: The impact of horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF ensemble prediction system. *Wea. Forecasting*, **17**, 173-191.
- Mylne, K. R., Evans, R. E., and Clark, R. T., 2002: Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Quart. J. Roy. Meteor. Soc.*, **128**, 361-384.
- Mylne, K.R., 2002: Decision-making from probability forecasts based on forecast value. *Meteorol. Appl.*, **9**, 307-315.
- Orrell, D, Smith, L. A., Barkmeijer, J., and Palmer, T. N., 2001: Model error in weather forecasting. *Nonlin. Proc. Geophys.*, **8**, 357-371.
- Orrell, D., 2002: Role of the metric in forecast error growth: how chaotic is the weather? *Tellus*, **54A**, 350-362
- Ott, E., B. R. Hunt, I. Szunyogh, M. Corazza, E. Kalnay, D.J. Patil, J. A. Yorke, A. V. Zimin, and E. J. Kostelich, 2003: Exploiting local low dimensionality of the atmospheric dynamics for efficient ensemble Kalman filtering. *Mon. Wea. Rev.*, submitted. Available from [eo4@umail.umd.edu](mailto:eo4@umail.umd.edu).

- Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.*, **127**, 279-304.
- Palmer, T. N., 2002: The economic value of ensemble forecasts as a tool for risk assessment: from days to decades. *Quart. J. Roy. Meteor. Soc.*, **128**, 747-774.
- Patil, D. J., B. R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, 2001: Local low dimensionality of atmospheric dynamics. *Phys. Rev. Letters*, **86**, 5878-5881.
- Pelly, J. L., and B. J. Hoskins, 2003: How well does the ECMWF ensemble prediction system predict blocking? *Quart. J. Roy. Meteor. Soc.*, **129**, 1683-1702.
- Pham, D. T., 2001 : Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, **129**, 1194-1207.
- Puri, K., Barkmeijer, J., and Palmer, T.N., 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.*, **127**, 709-731.
- Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002a: Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 103-114.
- Reichle, R. H., J. P. Walker, R. D. Koster, and P. R. Houser, 2002b: Extended vs. ensemble Kalman filtering for land data assimilation. *J. Hydromet.*, **3**, 728-740.
- Reynolds, C. A., and Palmer, T. N., 1999: Convergence of singular vectors toward Lyapunov vectors. *Mon. Wea. Rev.*, **127**, 2309-2323.
- Reynolds, C. A., and T. E. Rosmond, 2003: Nonlinear growth of singular-vector based perturbations. *Quart. J. Roy. Meteor. Soc.*, **129**, 3059-3078.

- Richardson, D. S., 2001a: Ensembles using multiple models and analyses. *Quart. J. Roy. Meteor. Soc.*, **127**, 1847-1864.
- Richardson, D. S., 2001b: Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size. *Quart. J. Roy. Meteor. Soc.*, **127**, 2473-2489.
- Roulston, M. S., and L. A. Smith, 2002: Evaluating probabilistic forecasts using information theory. *Mon. Wea. Rev.*, **130**, 1653-1660.
- Roulston, M. S., and L. A. Smith, 2003: Combining dynamical and statistical ensembles. *Tellus*, **55A**, 16-30.
- Roulston, M. S., D. T. Kaplan, J. Hardenberg, and L. A. Smith, 2003: Using medium-range weather forecasts to improve the value of wind energy production. *Renewable Energy*, **28**, 585-602.
- Sanders, F., Mullen, S. L., Baumhefner, D. P. 2000: Ensemble simulations of explosive cyclogenesis at ranges of 2-5 days. *Mon. Wea. Rev.*, **128**, 2920-2934.
- Sardeshmukh, P. D., C. Penland, and M. Newman, 2001: Rossby waves in a stochastically fluctuating medium. In *Stochastic Climate Models*, ed. P. Imkeller and J.-S. von Storch, *Progress in Probability*, Birkhaueser, Basel, pp. 369-384.
- Simmons, A. J., and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **128**, 647-677.
- Smith, L. A., Ziehmann, C., and Fraedrich, K., 1999: Uncertainty dynamics and predictability in chaotic systems. *Quart. J. Roy. Meteor. Soc.*, **125**, 2855-2886.

- Smith, L. A., 2000: Disentangling uncertainty and error: on the predictability of nonlinear systems. In *Nonlinear Dynamics and Statistics*, ed. Alistair E. Mees. Birkhauser, 31-64.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663-1677.
- Stephenson, D. B., and F. J. Doblas-Reyes, 2000: Statistical methods for interpreting Monte-Carlo forecasts. *Tellus*, **52A**, 300-322.
- Szunyogh, I., and Toth, Z. 2002: The effect of increased horizontal resolution on the NCEP global ensemble mean forecasts. *Mon. Wea. Rev.*, **130**, 1125-1143.
- Thornes, J. E., and Stephenson, D. B., 2001: How to judge the quality and value of weather forecast products. *Meteorol. Apps.*, **8**, 307-314.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485-1490.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: the generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.
- Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297-3319.
- Toth, Z., I. Szunyogh, and E. Kalnay, and G. Iyengar, 1999: Comments on "notes on the appropriateness of ``bred modes" for generating initial perturbations used in ensemble prediction." *Tellus*, **51A**, 442-449.
- Toth, Z., Zhu, Y., and Marchok, T. 2001: The use of ensembles to identify forecasts with small and large uncertainty. *Wea. Forecasting*, **16**, 463-477.

- Trevisan, A., F. Pancotti, and F. Molteni, 2001: Ensemble prediction in a model with flow regimes. *Quart. J. Roy. Meteor. Soc.*, **127**, 343-358.
- Tribbia, J. J., and D. P. Baumhefner, 2003: Scale interactions and atmospheric predictability: an updated perspective. *Mon. Wea. Rev.*, in press.
- Vannitsem, S., and Toth, Z., 2002: Short-term dynamics of model errors. *J. Atmos. Sci.*, **59**, 2594-2604.
- Verlaan, M. and A. W. Heemink, 2001: Nonlinearity in data assimilation applications. A practical method for analysis. *Mon. Wea. Rev.*, **129**, 1578-1589.
- Wandishin, M. S., S. L. Mullen, D. J. Stensrud, and H. E. Brooks, 2001: Evaluation of a short-range multimodel ensemble system. *Mon. Wea. Rev.*, **129**, 729-747.
- Wang, X., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140-1158.
- Wang, X., and C. Bishop, 2004: Which is better, an ensemble of positive/negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev.*, in review.
- Watson, J. S., and Colucci, S. J., 2002: Evaluation of ensemble predictions of blocking in the NCEP global spectral model. *Mon. Wea. Rev.*, **130**, 3008-3021.
- Wei, M., and Z. Toth, 2003: A new measure of ensemble performance: perturbation versus error correlation analysis (PECA). *Mon. Wea. Rev.*, **131**, 1549-1565.
- Whitaker, J. S., and T. M. Hamill, 2001: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913-1924.
- Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, in press. Available from [jeffrey.s.whitaker@noaa.gov](mailto:jeffrey.s.whitaker@noaa.gov).

Wilks, D. S., 2001. A skill score based on economic value for probability forecasts.

*Meteor. Applic.*, **8**, 209-219.

Wilks, D. S., 2003: The minimum spanning tree (MST) histogram as a verification tool for multidimensional ensemble forecasts. *Mon. Wea. Rev.*, in press.

Zhang, F., C. Snyder, and J. Sun, 2003a: Impacts of initial estimate and observations on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, accepted.

Zhang, Z., and Krishnamurti, T. N., 1999: A perturbation method for hurricane ensemble predictions. *Mon. Wea. Rev.*, **127**, 447-469.

Zhu, Y., Toth, Z., Wobus, R., Richardson, D., Mylne, K. 2002: The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.*, **83**, 73-83.